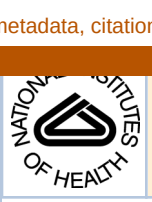

Published in final edited form as:

J Am Chem Soc. 2007 March 7; 129(9): 2660–2668. doi:10.1021/ja067971k.

Minimization of a Protein–DNA Dimerizer

Ryan L. Stafford[†], **Hans-Dieter Arndt**[†], **Mary L. Brezinski**[‡], **Aseem Z. Ansari**[‡], and **Peter B. Dervan**[*,†]

Contribution from the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, and Department of Biochemistry and the Genome Center, University of Wisconsins–Madison, Madison, Wisconsin 53706

Abstract

A protein–DNA dimerizer constructed from a DNA-binding polyamide and the peptide FYPWMKG facilitates the binding of a natural transcription factor Exd to an adjacent DNA site. The Exd binding domain can be reduced to a dipeptide WM attached to the polyamide through an ε-aminohexanoic acid linker with retention of protein–DNA dimerizer activity. Screening a library of analogues indicated that the tryptophan indole moiety is more important than methionine's side chain or the N-terminal acetamide. Remarkably, switching the stereochemistry of the tryptophan residue (L to D) stabilizes the dimerizer•Exd•DNA ternary complex at 37 °C. These observations provide design principles for artificial transcription factors that may function in concert with the cellular regulatory circuitry.

Introduction

Multivalent interactions are frequently encountered in biological systems.[1] Typically, the monovalent elements that mediate these molecular interactions use a small surface area and bind weakly. These monovalent elements are repeated, and the resulting multivalency improves the association between the interacting molecules. Dimerization is a key regulatory event in numerous biochemical settings, notably in signal transduction and the regulation of transcription.[2] Divalent small molecules called chemical inducers of dimerization or dimerizers have been used to facilitate protein–protein interactions.[2,3] Recently, we described a protein–DNA dimerizer, where one module binds a specific DNA sequence while the other interacts with a specific DNA-binding protein[4] (Figure 1). This molecule generates a bidentate surface that enhances the association of the targeted protein with its cognate DNA site. This divalent molecule improves the affinity of the targeted DNA-binding protein for its specific DNA sequence, whereas it is ineffective at sites where the DNA sequence does not match the preferred site of the target protein. Our long-term goal is to ask whether protein–DNA dimerizers could serve as artificial regulators of gene expression in a living cell.[4-6]

Transcription factors are modular in structure and combine DNA-binding domains with other functional domains and lead to the activation or repression of genes.1a Transcriptional activators themselves can be viewed as dimerizers composed of DNA and protein binding surfaces.3a Polyamides are ideal DNA-binding domains since they can be programmed to bind to a broad repertoire of DNA sequences using an aromatic amino acid pairing code.[8] A

[*] dervan@caltech.edu .
[†]California Institute of Technology.
[‡]University of Wisconsins– Madison.

hairpin polyamide–YPWM conjugate facilitates the binding a natural transcription factor extradenticle (Exd) to an adjacent DNA site.[4] The protein–DNA dimerizer **1** mimics the homeobox (HOX) transcription factor ultrabithorax (Ubx)[9a] (Figure 2a). Ubx cooperatively binds to DNA with the HOX cofactor protein Exd.[10] Their interaction is primarily mediated through Ubx's conserved YPWM peptide motif, which crosses the DNA minor groove to the major groove binding Exd.[9] Similarly, in human homologs as well, the YPWM motif is highly conserved within the HOX protein family and is often referred to as a hexapeptide ϕYPWMK, as the central YPWM tetrapeptide is often flanked by a hydrophobic residue (ϕ) and a charged residue, such as Lys or sometimes Arg.[11]

In this study, we address what are the minimum requirements for Exd recruitment. The Trp residue is strictly conserved, and a conservative mutation of Trp to Phe has been shown to lead to HOX proteins that no longer cooperatively bind to DNA with their cofactor proteins.[12] Met to Ile and Tyr to Leu mutations to the YPWM motif of HoxB4 have also been shown to eliminate cooperative binding.[13] The crystal structures of Ubx/Exd/DNA, HoxB1/Pbx1/DNA, and HoxA9/Pbx1/DNA complexes show the Trp indole deep in the protein binding pocket with an adjacent Met or Leu side chain reinforcing this interaction and mediating important protein contacts.[9] Thus, our efforts to minimize the protein-binding domain have centered on the Trp and Met residues.

The Exd-binding domain was reduced to the dipeptide WM and tethered to the polyamide DNA-binding domain through an ε-aminohexanoic acid linker. We report here that this conjugate with a smaller protein-binding domain retains significant dimerizing activity with Exd. A set of dipeptide analogues were screened to assess the importance of the remaining moieties. Minimized conjugates were then studied at physiologically relevant temperatures for fruit fly (20 °C) and human cell lines (37 °C).

Results

Protein–DNA Dimerizer•Exd Sequence Specificity

Exd is expected to bind 5′-TGAT-3′ as found in the Ubx/Exd/DNA crystal structure,[9a] and the hairpin polyamide ImImPyPy-γ-ImPyPyPy codes for the sequence 5′-WGWCCW-3′, where W equals A or T.[8] Exd (underlined) and polyamide sites were combined to create a composite site 5′-TGATTGACCA-3′ to study Exd recruitment by polyamide–FYPWMKG conjugate **1**[4] (Figure 2a). To confirm both the orientation preference of Exd recruitment as well as the specificity with respect to mismatch sites, MPE footprinting titrations[14] have been carried out on a DNA fragment featuring a polyamide mismatch site I (5′-TGATGGATTG-3′) and Exd upstream II (5′-TGATTGAC–CA-3′), as well as downstream III (5′-TGACCATGAT-3′) orientations (Figure 2b, right). An MPE footprint titration of dimerizer **1** without Exd protein present confirmed its highly specific binding, with an identical protection pattern for both sites II and III. Binding to site I, which is a single base pair mismatch for the polyamide DNA-binding domain in **1**, only occurred at ≥10 μM concentration, indicating a selectivity of better than 50-fold (Figure 2b, left). When 40 nM Exd was present with **1**, the MPE protection pattern was observed at a 3-fold lower concentration exclusively at site II, enhancing the selectivity over the site I to 150-fold and the Exd downstream site III to 3-fold. Furthermore, despite the presence of five sites for Exd on the DNA fragment, Exd alone at 1 μM concentration did not give rise to detectable binding. These results provide evidence for cooperative enhancement of binding of the dimerizer and Exd only when Exd's site is located upstream from the polyamide's. Subsequent experiments using DNA microarrays confirm that **1** prefers to recruit Exd to the consensus site 5′-NGANWGWC-3′ over other Exd–dimerizer DNA site spacings and orientations.[4c]

The Exd protein binding site in the presence of dimerizer **1** was determined to single base pair resolution (Figure 2b, middle). At low concentrations of Exd, dimerizer **1** occupied the polyamide site. With increasing Exd concentration, the MPE protection pattern became enlarged by four base pairs in the 3′-direction (^{32}P-labeled strand). Taking into account that an MPE footprint is generally observed two base pairs shifted to the 3′-end,[14] this maps protein binding to 5′-TGAT-3′ and dimerizer binding to the 5′-TGACCA-3′ (Figure 2c). Notably, no additional protein-induced protection was evident at mismatch site I or on the downstream arrangement site III. This observation supports the selectivity of the dimerization event, both for sequence composition and arrangement of the respective half-sites.

Analyzing Changes to the Protein-Binding Domain on Dimerization (4 °C)

The binding of Exd to DNA in the presence of various dimerizers was measured at 4 °C by electrophoretic mobility shift assays (EMSA). Exd protein binding was studied on a ^{32}P-labeled 47 base pair double-stranded DNA duplex containing a match composite binding site (5′-TGATTGACCA-3′) identical to site II in Figure 2b in the presence of compounds **1**–**5** and **15**–**22** (Figures 3 and 4). For comparison, a control titration with Exd alone was performed (top left, Figure 4). Negligible Exd binding was observed up to 100 nM concentrations; however, multiple Exd molecules appear to bind weakly and nonspecifically at 1 μM, in line with previous findings.[4] Thus, subsequent titrations were kept below this threshold and performed in the 10 pM to 100 nM range.

Upon the addition of 50 nM of conjugate **1**, the Exd DNA equilibrium association constant (K_a) was enhanced ($K_a = 1.5 \pm 0.3 \times 10^9$ M^{-1}),[15] yielding a nearly complete band shift (i.e., $\Theta_{app} \geq \sim 0.93 \pm 0.01$) at 5 nM Exd (Table 1). Thus, the presence of dimerizer **1** enhances Exd's DNA affinity by 1000-fold.[4] In controls, experiments with 50 nM of the parent polyamide **2**, which contains no peptide but has been shown previously by DNase footprinting experiments to bind specifically to the DNA-binding domain match site,[16] appears to enhance Exd's DNA affinity for DNA, as evidenced by the partial band shift (i.e., $\Theta_{app} = 0.5 \pm 0.2$).[17] The mismatch polyamide **3**, which is not expected to bind the polyamide match site,[18] does not lead to any measurable enhancement of Exd's DNA affinity. Thus, the DNA-binding polyamide bound by itself to the match site appears to facilitate Exd's binding to DNA but is insufficient for stable complex formation. As a further control, modified hairpin **4** was investigated to evaluate the effect of the proximal linker domain. The propylacetamide linker appeared to reduce the polyamide's contribution relative to hairpin **2**. Exd titrations in the presence of 50 nM of the YPWM peptide **5** did not lead to any measurable enhancement of Exd–DNA affinity, indicating that the polyamide DNA-binding domain is indispensable for dimerization.

The truncated dipeptide dimerizer **15** exhibited a similar binding affinity for DNA to **1** (see DNase I footprinting data below), so 50 nM concentration was used for the dipeptide analogues to saturate the DNA binding site and ensure direct comparisons could be made to changes to the presumed Exd binding domain. Polyamide–Ac–WM–dipeptide conjugate **15** enhanced Exd's DNA affinity almost 50% as effectively as the larger conjugate **1** ($K_a = 7 \pm 2 \times 10^8$ M^{-1}), with high levels of complex formation ($\Theta_{app} = 0.84 \pm 0.06$). When the stereochemistry of the Trp residue was inverted (**16**), Exd's DNA affinity remained the same ($K_a = 1.0 \pm 2 \times 10^9$ M^{-1}) with a slightly higher amount of complex formation ($\Theta_{app} = 0.91 \pm 0.03$). Substituting the Trp with Phe (**17**) or Ala (**18**) lowered Exd's DNA affinity so that it qualitatively appeared to be as effective as polyamide **2** by itself. In contrast to what was observed with the Trp residue, the Met residue could be converted to Ala (**19**) and Gly (**20**) without strongly affecting Exd's DNA affinity (K_a's = $6 \pm 1 \times 10^8$ M^{-1} and $5 \pm 2 \times 10^8$ M^{-1}, respectively) or the levels of complex formation ($\Theta_{app} > 0.75$). Noting that the stereochemistry of the Trp could be inverted (**15** to **16**), we examined whether the

stereogenic acetamide might be eliminated. Both conjugates **21** and **22** maintained similar Exd binding affinities (both K_a's = 7 ± 1 × 10^8 M^{-1}), albeit with slightly reduced levels of complex formation (Θ_{app} = 0.71 ± 0.10 and 0.65 ± 0.07, respectively). Interestingly, all of the minimal dimerizers that showed strong Exd recruitment yielded binding constants within 3-fold of conjugate **1**.

Analyzing Selected Dimerizers for Exd Binding at Higher Temperatures

The initial screen of dimerizers at 4 °C by gel shift yielded the promising minimal dimerizers **15** and **16**, which were further investigated (Figure 5 and Table 2). Specifically, the ability of conjugates **15** and **16** to recruit Exd at higher temperatures will be relevant to planned experiments in fruit flies (20 °C) and in human cell lines (37 °C).[19] Furthermore, the weak contribution of the DNA-binding domain by itself was investigated at these higher temperatures using compound **2**. The change from 4 to 20 °C did not significantly change Exd's K_a in the presence of compounds **1**, **15**, and **16** but markedly reduced the apparent effect seen by the polyamide DNA-binding domain **2**. Increasing the temperature further to 37 °C did not significantly affect Exd's K_a for compounds **1** and **16** but considerably impaired compound **15**'s ability to recruit Exd. Remarkably the only structural difference between compound **15** and **16** is the stereochemistry at the tryptophan residue, which apparently leads to a more stable ternary complex.

Conjugate DNA Binding Affinity and Sequence Specificity

The DNA affinity and sequence selectivity of conjugate **15** at 20 °C in the absence of Exd was determined by DNase I footprinting[14] (Figure 6 and Table 3). Binding energetics for compounds **1**–**3** have been reported previously.[4a,16,18] Experiments with **15** were performed using the same 3′-radiolabeled 250-base pair restriction fragment of a plasmid that was previously used for conjugate **1**. This fragment contains a 5′-TGGTCA-3′ match site IV as well as three single base pair mismatch sites I–III, 5′-AGCCA-3′, 5′-TGGGCA-3′, and 5′-TGGCCA-3′, respectively. The attachment of the glycine-linked Ac–FYPWMK–peptide to the DNA-binding domain of conjugate **1** was previously found to reduce the affinity approximately 100-fold with respect to the parent polyamide.[4a] Furthermore, the sequence specificity of **1** was diminished with respect to mismatch sites I and III, but it was still able to discriminate against mismatch site II by about 20-fold. The affinity of the minimized hairpin polyamide–WM (**15**) for its match site IV was improved by approximately 3-fold (K_a = 5.1 ± 1.0 × 10^8 M^{-1}). The sequence specificity was also found to be more than 25-fold for the match site IV with respect to mismatch sites I–III. Hence, the reduction in the size of the protein binding domain of the conjugate fully restored the intrinsic DNA sequence selectivity of the protein–DNA dimerizer.

Discussion

The observation that a dipeptide WM is sufficient for Exd recruitment is remarkable given that the YPWM motif is highly conserved with the consensus sequence ϕYPWMK (where ϕ denotes a hydrophobic residue) in the majority of HOX proteins[11] and that residues flanking the motif have been shown to contribute to the interaction.[12d] The total solvent-accessible surface area buried between Exd and the YPWM peptide of Ubx is ~570 Å[9a] where on average the buried surface area between more stable protein–protein interfaces is about 2000 Å.[20] From the crystal structures of the Ubx/Exd/DNA and HoxB1/Pbx1/DNA ternary complexes it was clear, however, that the majority of the contacts involved the Trp indole substituent, with the Met residue reinforcing the interaction.[9] Moreover, a related crystal structure of the HoxA9/Pbx1/DNA complex reveals a divergent AANWLH interaction motif bound in an entirely different conformation. The only distinct structural similarity between

the AANWLH motif of HoxA9 and the other YPWM motifs is the orientation of the Trp side chain in the protein binding pocket.

In line with these findings, the data presented here underscore the importance of the Trp side chain and to a lesser extent the Met residue. A “conservative” change of Trp to Phe, which reduces the hydrophobic interaction surface area and should eliminate the hydrogen bond to the main chain carbonyl of Leu23a (Exd homeodomain numbering),[9] abrogates the efficacy of the protein-binding domain of the dimerizer. Similarly, a Trp to Ala substitution is not tolerated. Truncating the Met side chain to a methyl or hydrogen only partially reduces efficacy but still provides a functional protein-binding domain. Inverting the stereochemistry of the Trp residue enhances stability, particularly at higher temperatures. It is likely that the indole moiety adopts a similar position in the Exd binding pocket as in the crystal structures. 9 Thus, the structural change enforced by the stereo-chemical switch (**15** and **16**) probably positions the acetamide in a favorable position for hydrogen-bonding contacts with Tyr25 of Exd, the DNA, or the dimerizer’s linker domain. Analogues lacking the acetamide retained some efficacy when compared to that of the L-Trp dimerizer **15** but do not lead to further insight into why the unnatural peptide of **16** is an improvement over **15**.

Our results support the notion that the geometry of the Exd protein and the protein-binding module in the DNA–dimerizer–protein complex do not deviate substantially from the structurally characterized ternary DNA–protein–protein complexes. Furthermore, MPE footprinting established Exd’s DNA binding site in the ternary DNA–dimerizer–protein complex as the expected 5′-TGAT-3′ match sequence upstream of the polyamide binding site 5′-TGACCA-3′. With the use of this information, a model of dimerizer **15** was constructed by superimposing a representative polyamide structure (PDB code 365D)21 on the Ubx/Exd/DNA ternary complex (PDB code 18BI).9a The DNA was replaced with an idealized B-form duplex (5′-AGGTGAT–TGACCACCAC-3′) created with 3DNA,22 and after making the requisite chemical changes, local energy minimization was performed23 (Figure 7). The three α-helix homeodomain of Exd is shown bound to the DNA major groove at 5′-TGAT-3′ adjacent to the minor groove-binding polyamide at 5′-TGACCA-3′ where the dimerizer protein-binding domain, which was previously part of Ubx, is attached by a linker to an internal pyrrole subunit in the hairpin.

Interestingly, the polyamide DNA-binding domain by itself was shown to contribute to the enhancement of Exd’s DNA affinity at low temperatures even though there are likely to be no direct contacts between the protein and the polyamide. The DNA minor groove is likely altered by the polyamide binding event, which may affect the adjacent major groove binding site for Exd. Polyamides have been shown to expand the width of the DNA minor groove by approximately 1 Å.[24] The minor groove polyamide binding site is also widened in the Exd/Ubx/DNA structure.[9a] The polyamide–DNA binding event appears to play an *allosteric* role in enhanced Exd binding. Undoubtedly, other changes cannot be ruled out, such as locally dampened structural dynamics, distortion of the nitrogenous bases, and reorientation of the phosphate backbone. We note that a seemingly innocuous propylacetamide linker on a pyrrole subunit appears to eliminate the polyamide’s weak energetic contribution, which underscores that small structural changes are at play. It is unclear if a simple dipeptide motif would be functional in the context of a natural protein such as Ubx.

In earlier studies we observed that the DNA binding selectivity for the polyamide–peptide conjugate **1** was reduced relative to that of the parent hairpin **2**.[4] The DNA specificity of the conjugate was restored by changing both the Trp and Met residues to Ala, which suggested that the Trp or Met might be responsible for the reduced recognition specificity.[4a] It is now clear, however, that a polyamide–peptide conjugate containing Trp and Met residues can exhibit sequence selectivity as evidenced by the footprinting of polyamide–WM conjugate

15. The diminished DNA binding affinity and specificity of polyamide–FYPWMKG conjugate **1** might be related to the overall size or hydrophobicity of the peptide, which suggests minimization may be a way to ensure specificity of DNA binding in the construction of polyamide conjugates and protein–DNA dimerizers in general.

Conclusion

Small molecule protein–DNA dimerizers allow the cooperative recruitment of the transcription factor Exd to a specific DNA sequence. Although the polyamide itself was found to contribute to Exd's DNA binding, the dipeptide WM protein binding domain was required for stable complex formation. Covalently linking polyamides to small molecule protein binders could emerge as a general strategy for the construction of artificial transcription factors.[25] The conjugate with the unnatural D-Trp is expected to reduce sensitivity to endogenous proteases and should function at biologically relevant temperatures. Although these minimized dimerizers might be further optimized, for instance through iterative changes to the linker length, composition of the polyamide, and various substitutions to the Trp indole, biologically driven projects utilizing the minimized dimerizers are in progress. Developmental processes in fruit flies have been perturbed using polyamides.[28] The question arises whether the polyamide–WM dimerizer can interfere specifically with the normal function of Exd leading to distinct fruit fly phenotypes[10a] which will be reported in due course.

Experimental Section

Materials

Wang resin and Fmoc-protected amino acids were purchased from Novabiochem unless stated otherwise. Super acid sensitive resin (SASRIN) was purchased from Bachem. Boc-*β*-Ala-PAM resin was purchased from Peptides International. Trifluoroacetic acid (TFA) was purchased from Halocarbon. Molecular biology grade ethylenediaminetetraacetic acid (EDTA), potassium glutamate, and acetylated bovine serum albumin (BSA) were purchased from Sigma. 3-Indolepropionic acid was purchased from Aldrich. *rac*-Dithiothreitol (DTT) was purchased from ICN. Other chemicals were purchased from Aldrich or EM Sciences and used without further purification. All other solvents were purchased from EM Sciences and were reagent grade. Water (18 MΩ) was purified using a Millipore MilliQ water purification system. Biological experiments were performed using Ultrapure distilled water (DNase/RNase free) purchased from Gibco. The pH of buffers was adjusted using a Beckman 340 pH/temp meter. All buffers were sterilized by filtration through either a Nalgene 0.2 μm cellulose nitrate filtration device or a Pall Acrodisc syringe filter HT Tuffryn membrane (0.2 μm). DNA oligonucleotides were obtained from the Biopolymer Synthesis and Analysis Facility at the Beckman Institute, California Institute of Technology. [α-^{32}P]-thymidine-5′-triphosphate (≥3000 Ci/mmol) and [α-^{32}P]-deoxyadenosine-5′-triphosphate (≥6000 Ci/mmol) were purchased from DuPont/NEN. [γ-^{32}P]-adenosine-5′-triphosphate (≥6000 Ci/mmol) was obtained from ICN. Calf thymus DNA was from Amersham, and all enzymes were obtained from Roche.

Methods

UV spectra were recorded in water using a Beckman Coulter DU 7400 spectrophotometer. All polyamide concentrations were determined using an extinction coefficient of ε = 69 500 $M^{-1}\cdot cm^{-1}$ at a λ_{max} near 310 nm. Peptide concentrations were determined using an extinction coefficient of ε = 5500 $M^{-1}\cdot cm^{-1}$ at λ_{max} equal to 280 nm. Matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed using an Applied Biosystems Voyager DE Pro spectrometer. Electrospray

ionization (ESI) mass spectrometry was performed using a Finnigan LCQ ion trap mass spectrometer. Analytical high-pressure liquid chromatography (HPLC) was performed with a Beckman Gold system equipped with a diode array detector using a Varian Microsorb-MV 300 C_{18} column (8 μm particle size, 250 mm × 4.6 mm). Preparative HPLC was performed with a Beckman Gold system equipped with a single-wavelength detector monitoring at 310 nm using a Waters DeltaPak C_{18} reversed-phase column (100 μm, 25 mm × 100 mm). For both analytical and preparative HPLC, solvent A was 0.1% (v/v) aqueous TFA and solvent B was acetonitrile. Compound purity was assessed using an HPLC analytical method employing a gradient of ~4.12% B/min starting from 10% B with a flow rate of 1.5 mL/min. Initial preparative HPLC was performed with a gradient of 10–60% B over 90 min. If the purity of the compound was less than 95%, a second HPLC purification was performed employing a slow gradient through the previous % B elution of the compound.

Synthesis

Polyamide monomers were prepared as described previously.[29,30] Conjugate **1** was synthesized as described previously.[4a] **1**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{108}H_{137}N_{32}O_{19}S^+$ 2218.1, observed 2218.0. Polyamides **2**, **3**, and **6** were synthesized on PAM resin using published BOC-based protocols and purified by reversed-phase HPLC.[29] **2**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{57}H_{71}N_{22}O_{10}^+$ 1223.6, observed 1223.6. Polyamide **3**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{58}H_{72}N_{21}O_{10}^+$ 1222.6, observed 1222.6. Polyamide **4**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{61}H_{78}N_{23}O_{11}^+$ 1308.6, observed 1308.7. Peptide **5** was synthesized by coupling 2 equiv of propanolamine with 1.1 equiv of 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and 1.1 equiv of *N*,*N*-diisopropylethylamine (DIEA) in *N*,*N*-dimethylformamide (DMF) to a previously reported protected peptide[4a] followed by deprotection and purification by reversed-phase HPLC as described. Peptide **5**: (ESI, positive) $[M + H]^+$ calcd for $C_{52}H_{71}N_{10}O_{10}S^+$ 1027.5, observed 1027.5. Polyamide **6** (R = H): (MALDI-TOF) calcd for $C_{59}H_{76}N_{23}O_{10}^+$ 1266.6, observed 1266.5. The peptides **7**–**12** were synthesized by manual solid-phase synthesis on Wang resin using standard Fmoc chemistry.[31] **7**: (ESI, negative) $[M–H]^-$ calcd for $C_{24}H_{33}N_4O_5S^-$ 489.2, observed 489.1. **8**: (ESI, negative) $[M–H]^-$ calcd for $C_{24}H_{33}N_4O_5S^-$ 489.2, observed 489.1. **9**: (ESI, negative) $[M–H]^-$ calcd for $C_{22}H_{32}N_3O_5S^-$ 450.2, observed 450.1. **10**: (ESI, negative) $[M–H]^-$ calcd for $C_{16}H_{28}N_3O_5S^-$ 374.2, observed 374.1. **11**: (ESI, negative) $[M–H]^-$ calcd for $C_{22}H_{29}N_4O_5^-$ 429.2, observed 429.1. **12**: (ESI, negative) $[M–H]^-$ calcd for $C_{21}H_{27}N_4O_5^-$ 415.2, observed 415.1. The peptides **13**–**14** were synthesized by manual solid-phase synthesis on SASRIN resin using standard Fmoc chemistry acylating the N-terminal residue with 3-indolepropionic acid activated with HBTU in situ. **13**: (ESI, negative) $[M–H]^-$ exact mass calcd for $C_{22}H_{30}N_3O_4S^-$ 432.2, observed 432.1. **14**: (ESI, negative) $[M–H]^-$ exact mass calcd for $C_{19}H_{24}N_3O_4^-$ 358.2, observed 358.1. Following cleavage from resin, crude peptides **7**–**14** were activated with HBTU and coupled to polyamide **6** (R = H) to generate polyamide conjugates **15**–**22**, respectively. Then conjugates **15**–**22** were purified by reverse phase HPLC to ≥95% purity (UV). Mass spectra of **15**, **16**, **17**, **18**, and **21** displayed the presence of small amounts (<5%) of methionine sulfoxide. **15**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{83}H_{108}N_{27}O_{14}S^+$ 1738.8, observed 1738.9. **16**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{83}H_{108}N_{27}O_{14}S^+$ 1738.8, observed 1738.8. **17**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{81}H_{107}N_{26}O_{14}S^+$ 1699.8, observed 1699.9. **18**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{75}H_{103}N_{26}O_{14}S^+$ 1623.8, observed 1623.9. **19**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{81}H_{104}N_{27}O_{14}^+$ 1678.8, observed 1678.9. **20**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{80}H_{102}N_{27}O_{14}^+$ 1664.8, observed 1664.9. **21**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{81}H_{105}N_{26}O_{13}S^+$ 1681.8, observed 1681.8. **22**: (MALDI-TOF) $[M + H]^+$ calcd for $C_{78}H_{99}N_{26}O_{13}^+$ 1607.8, observed 1607.8.

Protein Expression and Characterization

The *Drosophila melanogaster* extradenticle (Exd) homeodomain with its fourth extended helix (residues 238–320)[9] was expressed and FPLC-purified following the precedent of Passner et al. and stored in 25% glycerol at −20 °C.[9] The protein domain identity was confirmed by selective protease cleavage (Lys-C achromobacter peptidase) and MALDI-TOF MS, identifying 8 of 9 predicted fragments. MALDI-TOF MS of the undigested protein fragment was consistent with the isolation of Exd residues 238–320 (calculated mass, 9495.7 *m*/*Z*; observed mass, 9496.0 *m*/*Z*). Protein concentration was determined by measuring the UV absorption at 280 nm using a calculated[32] extinction coefficient of ε = 12 600 $M^{-1}\cdot cm^{-1}$.

Preparation of 3′-Labeled DNA for DNase I Footprinting

Plasmids pHDA1 or pDEH9 (5 μg) were digested using restriction enzymes *Pvu*II and *Eco*RI to yield the crude 250 base pair or 308 base pair restriction fragments.[4a] The 3′-overhang regions were filled with [α-^{32}P]-dATP and [α-^{32}P]-dTTP using the Klenow fragment of DNA polymerase or Sequenase. The radiolabeled oligonucleotides were purified by band extraction from a preparative 6% nondenaturing polyacrylamide gel, diluted to 10 kcpm/μL, stored at −78 °C, and used within 2 weeks.

Preparation of 5′ Labeled DNA for MPE Footprinting

A 5′-^{32}P-end-labeled (see below) primer 5′-ttc aca cag gaa aca gct atg acc-3′ and an unlabeled primer 5′-cgg gga tcc ata cat gat tt-3′ were used to amplify a 5′-lower strand-labeled PCR fragment from pHDA1 following published procedures.[14] The radiolabeled oligonucleotides were purified by band extraction from a preparative 6% nondenaturing polyacrylamide gel, diluted to 10 kcpm/μL, stored at −78 °C, and used within 2 weeks.

Preparation of ^{32}P-Labeled Synthetic Oligonucleotides for EMSA

Synthetic oligonucleotides were either labeled by (a) 3′-overhang fillin or (b) 5′ kinase end labeling. Three synthetic oligonucleotides were used: (1) UExdOpt: 5′-gatctcccgg cgaatgattg accatatcgg cgccactgtc acccgga-3′, (2) LExdOpt: 5′-tccgggtgac agtggcgccg atatggtcaa tcattcgccg ggagatc-3′, and (3) LSExdOpt: 5′-tccgggtgac agtggcgccg atatggtcaa tcattcgccg ggag-3′. Annealing was accomplished by mixing equimolar quantities in water, heating to 95 °C for 5 min, and slowly cooling to room temperature over 2 h. The 3′ fill-in was accomplished by annealing UExdOpt and LSExdOpt DNA (~60 pmol) followed by treatment with [α-^{32}P]-ATP and [α-^{32}P]-dTTP and the Klenow fragment of DNA polymerase. The 5′ end labeling was accomplished by annealing UExdOpt and LExdOpt followed by labeling with [γ-^{32}P]-ATP and polynucleotide kinase using standard protocols. Both 3′- and 5′-labeled oligonucleotides were then purified by gel electrophoresis and band extraction from a 6% nondenaturing polyacrylamide gel, diluted to 10 kcpm/μL, stored at −78 °C, and used within 2 weeks.

Electrophoretic Mobility Shift Assay (EMSA)

EMSA experiments were performed as described previously.[4a] Polyamide conjugates were preincubated with ≤5 kcpm/reaction of ^{32}P-labeled oligonucleotides at room temperature (~20 °C) in buffer for 1 h. Exd protein was then added to the reaction solution in a cold room (~4 °C), at room temperature (~20 °C), or in an incubator (~37 °C) for 1 more hour. The final concentrations were 30 mM potassium glutamate, 10 mM *N*-[2-Hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid] (HEPES), 0.2 mM DTT, 10% glycerol, and 0.1 mg/mL BSA (Sigma). A volume of 10 μL of the binding solution was loaded onto a nondenaturing 9% polyacrylamide (29:1 acrylamide/bisacrylamide)/3% glycerol/1X TBE (89 mM trishydroxymethylaminomethane (TRIS) base, 89 mM boric acid, 2 mM EDTA, pH

8.3) 0.8 mm thick gel and developed at 180–190 V and 13–15 mA for 1.5 h pre-equilibrated to the same temperature in which the final Exd incubation was performed (i.e., 4, 20, or 37 °C). Gels were then dried and exposed to a phosphor screen overnight, which were scanned with a Molecular Dynamics 400S phosphorimager. All experiments were done at least in triplicate.

The amount of Exd bound to the DNA template, i.e., the apparent fractional occupancy (Θ_{app}), was calculated from the measured intensity volumes of the shifted ($V_{shifted}$) and the unshifted bands ($V_{unshifted}$):

$$\Theta_{app} = V_{shifted} / (V_{shifted} + V_{unshifted}) \quad (1)$$

The data were then fit to the Hill equation:[33]

$$\Theta_{fit} = \Theta_{min} + (\Theta_{max} - \Theta_{min}) [(K_a{}^n [Exd]^n) / (1 + K_a{}^n [Exd]^n)] \quad (2)$$

where [Exd] is the Exd concentration, K_a is the equilibrium association constant, n is the Hill constant, and Θ_{min} and Θ_{max} are the experimentally determined Θ_{app} values where binding is absent or maximally saturated, respectively. KaleidaGraph software was used to fit the data to eq 2 using a least-squares fitting procedure minimizing the difference between Θ_{app} and Θ_{fit} with K_a, Θ_{max}, and Θ_{min} as the variable parameters. Fitting the data to eq 2 with free n did not increase the quality ($R \geq 0.98$ in either case), so $n = 1$ was kept for all analysis.

Affinity Determination by Quantitative DNase I Footprinting

Reactions were carried out in a volume of 400 μL in aqueous TKMC buffer according to published protocols[14] using a 3′ ^{32}P-labeled 250 base pair *Eco*RI/*Pvu*II restriction fragment of the plasmid pDEH9.[4a] Developed gels were imaged using storage phosphor autoradiography using a Molecular Dynamics 400S phosphorimager. Equilibrium association constants were determined as previously described.[14]

MPE Footprinting in the Presence of Exd Protein

MPE foot-printing was carried out in 40 μL reaction volumes according to published procedures.[14] Reactions were carried out in a total volume of 40 μL in aqueous TKMC buffer containing 10% glycerol using the 3′ ^{32}P-labeled 308 base pair *Eco*RI/*Pvu*II restriction fragment of the plasmid pHDA1. DNA was preincubated with dimerizer overnight at 20 °C: 4 μL of glycerol, 8 μL of 5× TKMC, 22 μL of DNA (20 kcpm), 4 μL of dimerizer (10×). Exd protein (0.16 μM, 1 μL) was introduced, gently mixed, and the mixture was incubated for 30 min at 4 °C before digestion was initiated. Final concentrations were 25 mM TrisOAc, 10 mM NaCl, 100 μM base pair carrier DNA (calf thymus), 10 mM DTT, 10 μM MPE × $FeSO_4$, 10% glycerol, pH 7.0. Cleavage reactions were stopped after 10 min.

Acknowledgments

This work was supported by the National Institute of Health (Research Grants to P.B.D. and A.Z.A). A DAAD postdoctoral fellowship (to H.D.A) is gratefully acknowledged. We thank Professor A. L. Aggarwal (Mount Sinai School of Medicine, New York) for a donation of the plasmid encoding the Exd DNA binding domain and Drs. P. Snow and G. Hathaway (Beckman Institute, California Institute of Technology) for technical assistance with protein expression, purification, and identification.

References

(1). (a) Ptashne, M.; Gann, A. Genes & Signals. Cold Spring Harbor Laboratory Press; Plainview, NY: 2002. (b) Mammen M, Choi S-K, Whitesides GM. Angew. Chem., Int. Ed. 1998; 37:2754–2794. (c) Kiessling LL, Gestwicki JE, Strong LE. Curr. Opin. Chem. Biol. 2000; 4:696–703. [PubMed: 11102876] (d) Harmer NJ, Chirgadze D, Hyun Kim K, Pellegrini L, Blundell TL. Biophys. Chem. 2003; 100:545–553. [PubMed: 12646390]

(2). Klemm JD, Schreiber SL, Crabtree GR. Annu. Rev. Immunol. 1998; 16:569–592. [PubMed: 9597142]

(3). (a) Austin DJ, Crabtree GR, Schreiber SL. Curr. Biol. 1994; 1:131–136. (b) Farrar MA, Alberola-Ila J, Perlmutter RM. Nature. 1996; 383:178–181. [PubMed: 8774884] (c) Choi J, Chen J, Schreiber SL, Clardy J. Science. 1996; 273:239–242. [PubMed: 8662507] (d) Schreiber SL. Bioorg. Med. Chem. 1998; 6:1127–1152. [PubMed: 9784856] (e) Althoff EA, Cornish VW. Angew. Chem. 2002; 114:2433–2436.Angew. Chem., Int. Ed. 2002; 41:2327–2330.

(4). (a) Arndt HD, Hauschild KE, Sullivan DP, Lake K, Dervan PB, Ansari AZ. J. Am. Chem. Soc. 2003; 125:13322–13323. [PubMed: 14583004] (b) Hauschild KE, Metzler RE, Arndt H-D, Moretti R, Raffaelle M, Dervan PB, Ansari AZ. Proc. Natl. Acad. Sci. U.S.A. 2005; 102:5008–5013. [PubMed: 15781856] (c) Warren CL, Kratochvil NCS, Hauschild KE, Foister S, Brezinski ML, Dervan PB, Phillips GN Jr. Ansari AZ. Proc. Natl. Acad. Sci. U.S.A. 2006; 103:867–872. [PubMed: 16418267]

(5). (a) Denison C, Kodadek T. Chem. Biol. 1998; 5:R129–R145. [PubMed: 9653545] (b) Ansari AZ, Mapp AK. Curr. Opin. Chem. Biol. 2002; 6:765–772. [PubMed: 12470729] (c) Arndt H-D. Angew. Chem., Int. Ed. 2006; 45:4552–4560.

(6). Protein–protein dimerizers have been employed in engineered systems as a means for exercising external control of gene expression of exogenously introduced genes (ref 7).

(7). (a) Pollock R, Clackson T. Curr. Opin. Biotechnol. 2002; 13:459–467. [PubMed: 12459338] (b) Minter AR, Brennan BB, Mapp AK. J. Am. Chem. Soc. 2004; 126:10504–10505. [PubMed: 15327284] (c) Liu B, Alluri PG, Yu P, Kodadek T. J. Am. Chem. Soc. 2005; 127:8254–8255. [PubMed: 15941237]

(8). (a) Dervan PB. Bioorg. Med. Chem. 2001; 9:2215–2235. [PubMed: 11553460] (b) Dervan PB, Edelson BS. Curr. Opin. Struct. Biol. 2003; 13:284–299. [PubMed: 12831879]

(9). (a) Passner JM, Ryoo HD, Shen LY, Mann RS, Aggarwal AK. Nature. 1999; 397:714–719. [PubMed: 10067897] (b) Piper DE, Batchelor AH, Chang CP, Cleary ML, Wolberger C. Cell. 1999; 96:587–597. [PubMed: 10052460] (c) LaRronde-LeBlanc NA, Wolberger C. Genes Dev. 2003; 17:2060–2072. [PubMed: 12923056]

(10). (a) Peifer M, Wieschaus E. Genes Dev. 1990; 4:1209–1223. [PubMed: 1976570] (b) Chan SK, Jaffe L, Capovilla M, Botas J, Mann RS. Cell. 1994; 78:603–615. [PubMed: 7915199] (c) van Dijk MA, Murre C. Cell. 1994; 78:617–624. [PubMed: 7915200] (d) Mann RS, Chan SK. Trends Genet. 1996; 12:258–262. [PubMed: 8763497]

(11). Duboule, D., editor. Guidebook to the Homeobox Genes. Oxford University Press; Oxford: 1994.

(12). (a) Neuteboom STC, Peltenburg LTC, van Dijk MA, Murre C. Proc. Natl. Acad. Sci. U.S.A. 1995; 92:9166–9170. [PubMed: 7568094] (b) Knoepfler PS, Kamps MP. Mol. Cell. Biol. 1995; 15:5811–5819. [PubMed: 7565734] (c) Lu Q, Kamps MP. Mol. Cell. Biol. 1996; 16:1632–1640. [PubMed: 8657138] (d) Shanmugam K, Featherstone MS, Saragovi HU. J. Biol. Chem. 1997; 272:19081–19087. [PubMed: 9228093]

(13). Shen WF, Chang CP, Rozenfeld S, Sauvageau G, Humphries RK, Lu M, Lawrence HJ, Cleary ML, Largman C. Nucleic Acids Res. 1996; 24:898–906. [PubMed: 8600458]

(14). Trauger JW, Dervan PB. Methods Enzymol. 2001; 340:450–466. [PubMed: 11494863]

(15). These affinities lie within 2-fold of those reported previously (ref 4a). The measured difference in binding affinity may be due to variation of Exd activities when using different protein preparations.

(16). Marques MA, Doss RM, Urbach AR, Dervan PB. Helv. Chim. Acta. 2002; 85:4485–4517.

(17). Although there are reports of “caging” effects in gel shifts (i.e., protein–DNA complexes with dissociation half-lives that are shorter than the time it takes to run a gel often form discrete

supershifted bands) the experiments reported here have not been optimized to stabilize weak complexes. Under the conditions optimized for our initial dimerizer (**1**) the complex formed in the presence of **2** is not stable; thus, we do not make a definitive K_a assignment.

(18). Gottesfeld JM, Melander C, Suto RK, Raviol H, Luger K, Dervan PB. J. Mol. Biol. 2001; 309:615–629. [PubMed: 11397084]

(19). In earlier studies, we investigated the role of the linker that connects the polyamide to the peptide (ref 4b). We find that a substantial degree of variability in the linker length is tolerated at lower temperatures. At physiological temperatures, the longest linker tested confers a “switchlike” property on the protein–DNA dimerizer, in that it abolishes the ability of the YPWM moiety to recruit the natural transcription factor to DNA.

(20). (a) Chakrabarti P, Janin J. Proteins. 2002; 47:334–343. [PubMed: 11948787] (b) Lo Conte L, Chothia C, Janin J. J. Mol. Biol. 1999; 285:2177–2198. [PubMed: 9925793]

(21). Kielkopf CL, Baird EE, Dervan PD, Rees DC. Nat. Struct. Biol. 1998; 5:104–109. [PubMed: 9461074]

(22). Lu XJ, Olson WK. Nucleic Acids Res. 2003; 31:5108–5121. [PubMed: 12930962]

(23). The illustrative models shown in Figure 7 were created from the atomic coordinates for Exd (PDB code 18BI) and the polyamide (PDB code 365D) derived from the crystal structures and superimposed with a “Fit” algorithm on the DNA backbone using SwissPDB Viewer (ver. 3.7). The DNA coordinates were derived from 3DNA software using the Fiber program with ideal B-form DNA parameters and superimposed as above. The DNA and Ubx, except for the WM motif, from the previous structures were then deleted with SwissPDB Viewer. The linker between the polyamide and the WM, the γ-aminobutyric acid of the polyamide hairpin, and the N-terminal acetyl moiety were constructed and minimized using an augmented MM3 force field with all other coordinates locked using CAChe Workspace (ver. 6.1.10). Finally, the minor groove width was relaxed at the polyamide binding site through a short (<10 ps) MD simulation in CAChe on the bases GGTC and CCAG with all other coordinates locked.

(24). (a) Kielkopf CL, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC. Science. 1998; 282:111–115. [PubMed: 9756473] (b) Kielkopf CL, Bremer RE, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC. J. Mol. Biol. 2000; 295:557–567. [PubMed: 10623546] (c) Suto RK, Edayathumangalam RS, White CL, Melander C, Gottesfeld JM, Dervan PB, Luger K. J. Mol. Biol. 2003; 326:371–380. [PubMed: 12559907] (d) Zhang Q, Dwyer TJ, Tsui V, Case DA, Cho JH, Dervan PB, Wemmer DE. J. Am. Chem. Soc. 2004; 126:7958–7966. [PubMed: 15212545] (e) Edayathumangalam RS, Weyermann P, Gottesfeld JM, Dervan PB, Luger K. Proc. Natl. Acad. Sci. U.S.A. 2004; 101:6864–6869. [PubMed: 15100411]

(25). In a formal sense, dimerizers equipped with DNA and protein-binding surfaces could induce specific proximal interactions between DNA and proteins other than transcription factors, such as coactivators (ref 26) and DNA-modifying enzymes (ref 27).

(26). (a) Kuznetsova S, Ait-Si-Ali S, Nagibneva I, Troalen F, Le Villain J, Harel-Bellan A, Svinarchuk F. Nucleic Acids Res. 1999; 27:3995–4000. [PubMed: 10497263] (b) Mapp AK, Ansari AZ, Ptashne M, Dervan PB. Proc. Natl. Acad. Sci. U.S.A. 2000; 97:3930–3935. [PubMed: 10760265] (c) Ansari AZ, Mapp AK, Nguyen DH, Dervan PB, Ptashne M. Chem. Biol. 2001; 8:583–592. [PubMed: 11410377] (d) Liu B, Han Y, Corey DR, Kodadek T. J. Am. Chem. Soc. 2002; 124:1838–1839. [PubMed: 11866581] (e) Stanojevic D, Young RA. Biochemistry. 2002; 41:7209–7216. [PubMed: 12044151] (f) Arora PS, Ansari AZ, Best TP, Ptashne M, Dervan PB. J. Am. Chem. Soc. 2002; 124:13067–13071. [PubMed: 12405833] (g) Liu B, Han Y, Ferdous A, Corey DR, Kodadek T. Chem. Biol. 2003; 10:909–916. [PubMed: 14583257] (h) Kwon Y, Arndt H-D, Mao Q, Choi Y, Kawazoe Y, Dervan PB, Uesugi M. J. Am. Chem. Soc. 2004; 126:15940–15941. [PubMed: 15584709]

(27). (a) Matteucci M, Lin K-Y, Huang T, Wagner R, Sternbach DD, Mehrotra M, Besterman JM. J. Am. Chem. Soc. 1997; 119:6939–6940. (b) Wang CC, Dervan PB. J. Am. Chem. Soc. 2001; 123:8657–8661. [PubMed: 11535069] (c) Arimondo PB, Bailly C, Boutorine AS, Ryabinin VA, Syniakov AN, Sun J-S, Garestier T, Helene C. Angew. Chem., Int. Ed. 2001; 40:3045–3048.

(28). (a) Janssen S, Durussel T, Laemmli UK. Mol. Cell. 2000; 6:999–1011. [PubMed: 11106740] (b) Janssen S, Cuvier O, Muller M, Laemmli UK. Mol. Cell. 2000; 6:1013–1024. [PubMed: 11106741]

(29). Baird EE, Dervan PB. J. Am. Chem. Soc. 1996; 118:6141–6146.

(30). Rucker VC, Foister S, Melander C, Dervan PB. J. Am. Chem. Soc. 2003; 125:1195–1202. [PubMed: 12553822]

(31). Wellings DA, Atherton E. Methods Enzymol. 1997; 289:44–67. [PubMed: 9353717]

(32). Gill SC, von Hippel PH. Anal. Biochem. 1989; 182:319–326. [PubMed: 2610349]

(33). Cantor, CR.; Schimmel, PR. Biophysical Chemistry of Macromolecules, Part 3: The BehaVior of Biological Macromolecules. W.H. Freeman and Company; New York: 1980.

(34). Delano, WL. The PyMOL Molecular Graphics System. DeLano Scientific; San Carlos, CA: 2002. http://www.pymol.org

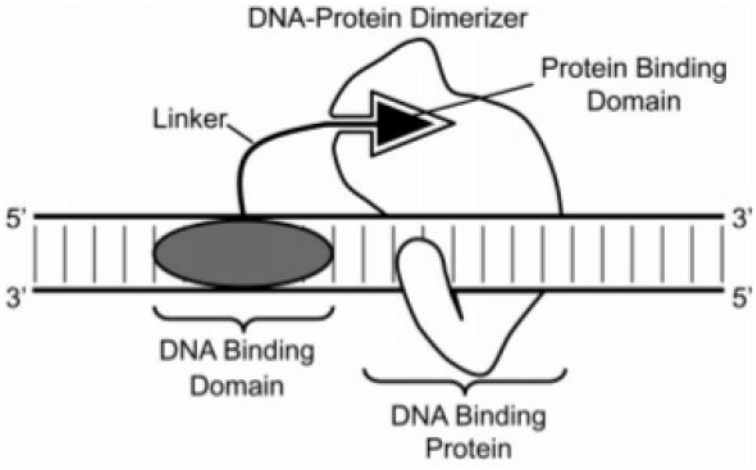

Figure 1.
Protein–DNA dimerizer facilitating the binding of a protein to an adjacent DNA site.

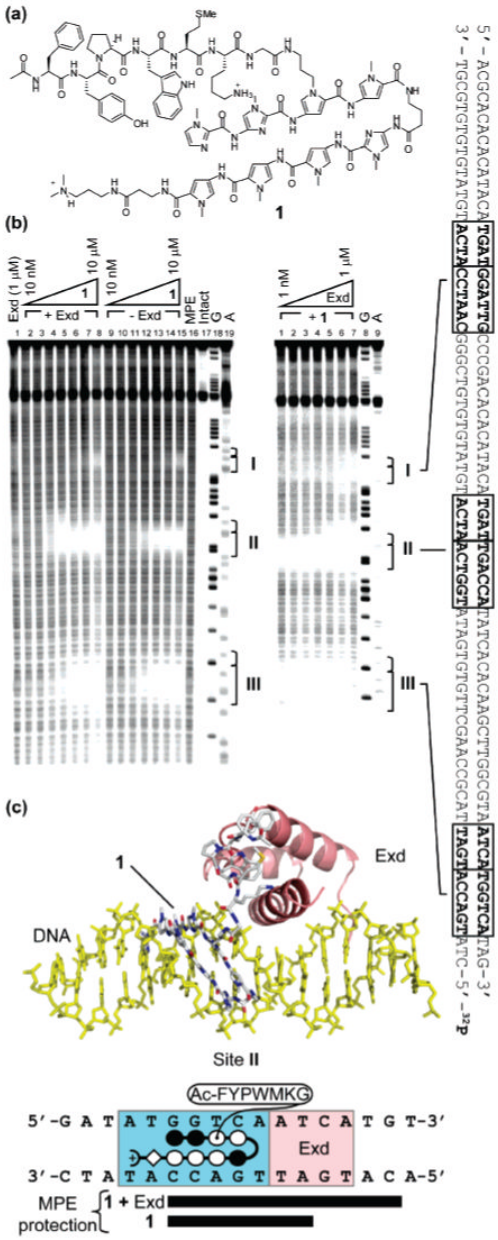

Figure 2.

(a) Dimerizer **1** structure. (b) (left) MPE footprint titration of conjugate **1** on a fragment of pHDA1 in the presence (40 nM, lanes 2–8) and absence of Exd protein (lanes 9–15) and 100 μM base pair calf thymus DNA. Lane 1: Exd only (1 μM). Lanes 2–8: 0.01, 0.03, 0.1, 0.3, 1, 3, 10 μM conjugate **1** all with 40 nM Exd. Lanes 9–15: 0.01, 0.03, 0.1, 0.3, 1, 3, 10 μM conjugate **1**. Lane 16: MPE standard. Lane 17: Intact DNA. Lane 18: G-sequencing lane. Lane 19: A-sequencing lane. (middle) MPE footprint titration of Exd in the presence of dimerizer **1** and 100 μM base pair calf thymus DNA. Lanes 1–7: 0.01, 0.03, 0.1, 0.3, 1, 3, 10 μM Exd all with 1 μM conjugate **1**. Lane 8: G-sequencing lane. Lane 9: A-sequencing lane. (right) Sequence of fragment of pHDA1 used for MPE footprinting. (c) Structural model of dimerizer **1** in a complex with DNA and Exd. A key is shown below with a dimerizer ball-and-stick model which indicates where the MPE protection pattern was observed (filled

circle = *N*-methylimidazole, empty circle = *N*-methylpyrrole, diamond = β-Ala, halfcircle with plus = *N*,*N*-dimethylaminopropylamine, half-circle = *γ*-aminobutyric acid).

Figure 3.
Structures of compounds **2**–**22** and the synthesis of minimized polyamide–dipeptide conjugates **15**–**22**. i: **6** (R = H), HBTU, DMF, excess DIEA, room temperature, 1–3 h.

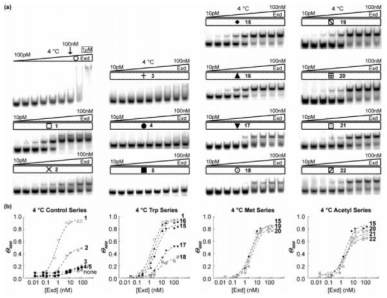

Figure 4.

Representative gel shift assays at 4 °C all done in triplicate. Autoradiograms are shown for Exd titrations with a 47-base pair ^{32}P-labeled dsDNA probe containing both Exd (underlined) and dimerizer binding sites (5′-TGATTGACCA-3′). (a) The upper left image shows increasing concentrations of Exd only from left to right at 100 pM, 500 pM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, 500 nM, and 1 μM. In all other experiments the indicated compound was applied at 50 nM, and increasing concentrations of Exd were studied from left to right: 10 pM, 50 pM, 100 pM, 500 pM, 1 nM, 5 nM, 10 nM, 50 nM, and 100 nM. The upper band consists of the Exd–dimerizer–DNA complex, while the lower band consists of the dimerizer–DNA complex. (b) Isotherms for Exd binding at 4 °C organized into groups.

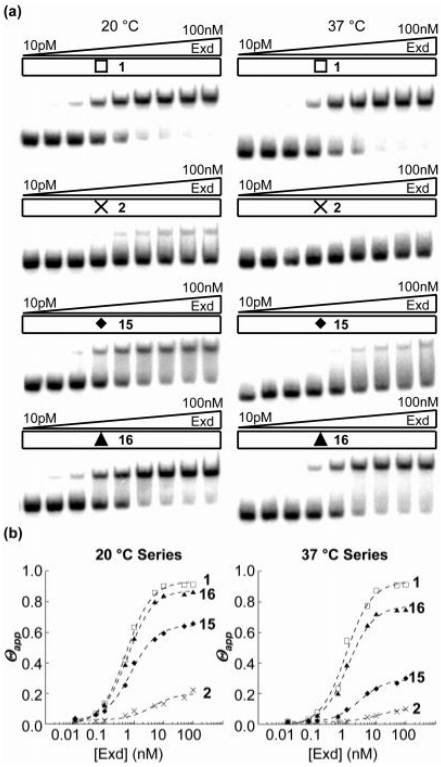

Figure 5.

(a) Representative gel shift assays at 20 and 37 °C all done in triplicate shown in the same format as in Figure 4. (b) Isotherms for Exd binding at 20 and 37 °C organized into groups.

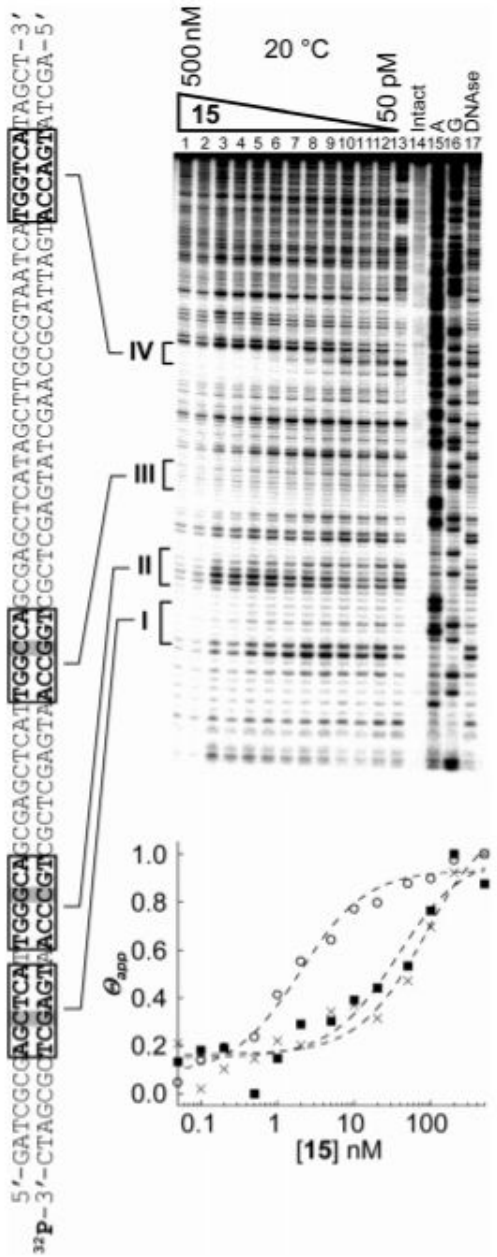

Figure 6.

Quantitative DNase I footprinting titration of minimized dimerizer **15** on a restriction fragment containing a match site IV (5′-TGGTCA-3′) and three mismatch sites I–III, (5′-AG**C**TCA-3′), (5′-TGG**G**CA-3′), and (5′-TGG**C**CA-3′), respectively. The analyzed binding site locations are indicated by square brackets and correlated to the sequence on the left. Lane 1–13: 500, 200, 100, 50, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1, and 0.050 nM of compound **15**. Lane 14: Intact DNA. Lane 15: A-sequencing lane. Lane 16: G-sequencing lane. Lane 17: DNase I standard. The isotherms for binding to sites I (×), III (■), and IV (○) are shown below.

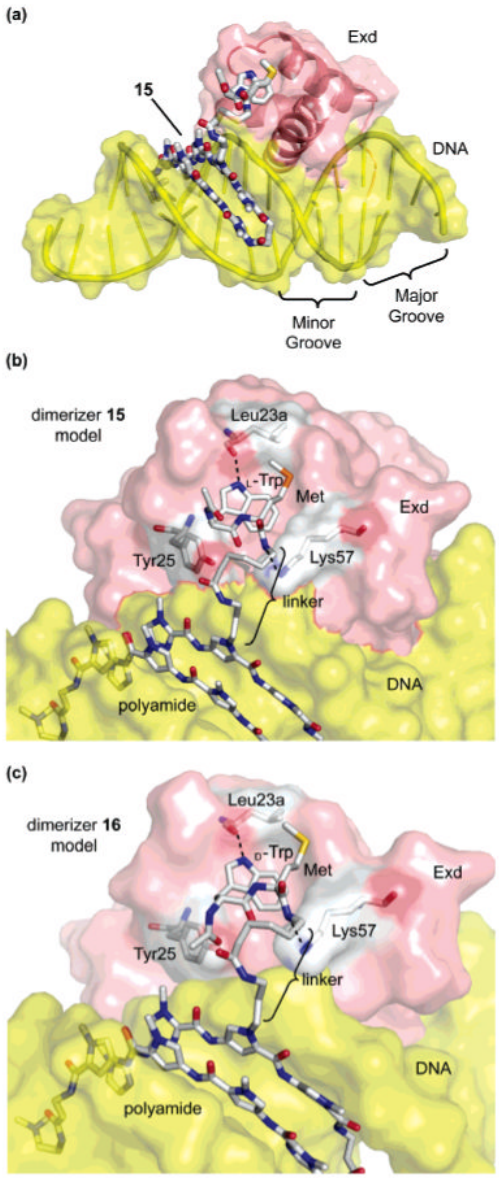

Figure 7.

(a) Model of the protein–DNA dimerizer **15** (sticks) bound to the minor groove of DNA (yellow cartoon and surface) adjacent to the recruited major groove-binding Exd homeodomain (pink cartoon and surface). Attached by a linker to a polyamide pyrrole residue, the Ac–WM–dipeptide protein-binding domain of **15** projects out of the DNA minor groove and interacts with the Exd binding pocket. (b) Closer view of the modeled interaction of **15** (sticks), Exd (transparent red surface), and DNA (transparent yellow surface) centered on the Exd binding pocket. Select Exd homeodomain residues are numbered (ref 9a) (white = carbon; dark blue = nitrogen; red = oxygen). Images were generated using PyMOL (ref 34). (c) Modeled interaction of **16**, Exd, and DNA with same color coding as in panel b.

Table 1

Exd–DNA–Dimerizer Stabilities at 4 °C

compd	K_a of Exd [10^9 M^{-1}][a]	Θ_{max}[b]
1	1.5 ± 0.3	0.93 ± 0.01
15	0.7 ± 0.2	0.84 ± 0.06
16	1.0 ± 0.2	0.91 ± 0.03
19	0.6 ± 0.1	0.84 ± 0.04
20	0.5 ± 0.2	0.77 ± 0.08
21	0.7 ± 0.1	0.71 ± 0.10
22	0.7 ± 0.1	0.65 ± 0.07

[a] Mean of ≥3 titrations of Exd at 50 nM compound. Data have been fit to a Hill equation with $n = 1$ ($R > 0.98$).

[b] Determined at least three times at 100 nm Exd and 50 nm dimerizer.

Table 2

Exd–DNA–Dimerizer Stabilities at 20 and 37 °C

compd	K_a of Exd [10^9 M^{-1}] at 20 °C[a]	Θ_{max} at 20 °C[b]	K_a of Exd [10^9 M^{-1}] at 37 °C[a]	Θ_{max} at 37 °C[b]
1	1.9 ± 0.2	0.91 ± 0.01	1.1 ± 0.2	0.91 ± 0.02
15	1.3 ± 0.1	0.65 ± 0.11	*c*	0.30 ± 0.12
16	1.7 ± 0.1	0.86 ± 0.02	0.9 ± 0.1	0.75 ± 0.06

[a] Mean of ≥3 titrations of Exd at 50 nM compound. Data have been fit to a Hill equation with $n = 1$ ($R > 0.99$).

[b] Determined at least three times at 100 nm Exd and 50 nm dimerizer.

[c] Not determined because no stable complex was detected (i.e., mean $\Theta_{max} \leq 0.5$).

Table 3

Conjugate–DNA Equilibrium Association Constants K_a's [10^8 M^{-1}][a]

compd	5′-TGGTCA-3′	5′-TGGCCA-3′	5′-TGGGCA-3′	5′-AGCTCA-3′
1 [b]	1.8 ± 0.2	1.3 ± 0.2 [1.4]	≤0.1 [> 18]	1.5 ± 0.3 [1.2]
15	5±1	0.2 ± 0.1 [25]	≤0.1 [> 50]	0.12 ± 0.05 [42]

[a] Specificity ratio with respect to match site IV (5′-TGGTCA-3′) shown in brackets.

[b] See ref 4a.